

18.700 LECTURE NOTES, 11/10/04

CONTENTS

1.	Direct sum decompositions	1
2.	Generalized eigenspaces	3
3.	The Chinese remainder theorem	5
4.	Linear independence of generalized eigenspaces	8

1. DIRECT SUM DECOMPOSITIONS

Definition 1.1. Let V be an \mathbb{F} -vector space. Let $\mathcal{W} = (W_1, \dots, W_s)$ be an ordered s -tuple of vector subspaces of V .

- (i) An *ordered s -tuple of vectors in \mathcal{W}* , $(\mathbf{w}_1, \dots, \mathbf{w}_s)$, is an ordered s -tuple of vectors in V such that for every $i = 1, \dots, s$, $\mathbf{w}_i \in W_i$.
- (ii) The ordered s -tuple \mathcal{W} is *linearly independent* if the only ordered s -tuple of vectors in \mathcal{W} , $(\mathbf{w}_1, \dots, \mathbf{w}_s)$, satisfying $\mathbf{w}_1 + \dots + \mathbf{w}_s = \mathbf{0}$ is $(\mathbf{0}, \dots, \mathbf{0})$.
- (iii) The ordered s -tuple \mathcal{W} *spans V* , or *is spanning*, if for every vector $\mathbf{v} \in V$, there exists an ordered s -tuple of vectors in \mathcal{W} , $(\mathbf{w}_1, \dots, \mathbf{w}_s)$, satisfying $\mathbf{w}_1 + \dots + \mathbf{w}_s = \mathbf{v}$.
- (iv) The ordered s -tuple \mathcal{W} is a *direct sum decomposition of V* if it is linearly independent and spans V .

Remark 1.2. Please do not confuse the notion of linearly independent (resp. spanning) collection of subspaces of V with the notion of linearly independent (resp. spanning) collection of vectors in V . These are closely related, but different.

Proposition 1.3. An ordered s -tuple of subspaces of V , $\mathcal{W} = (W_1, \dots, W_s)$, is linearly independent iff for every ordered s -tuple $(\mathcal{B}_1, \dots, \mathcal{B}_s)$ of linearly independent ordered subsets $\mathcal{B}_i \subset W_i$, the concatenation $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ is a linearly independent ordered subset of V .

Proof. (\Rightarrow) Assume \mathcal{W} is linearly independent. For every $i = 1, \dots, s$, define $e_i = \#\mathcal{B}_i$, which is 0 if \mathcal{B}_i is empty. Define $e = e_1 + \dots + e_s$. For every $i = 1, \dots, s$ such that $e_i > 0$, denote $\mathcal{B}_i = (\mathbf{w}_{(i,1)}, \dots, \mathbf{w}_{(i,e_i)})$. If $e = 0$, then $\mathcal{B} = \emptyset$. Otherwise, by definition, \mathcal{B} is the unique ordered set of vectors $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_e)$ such that for every $i = 1, \dots, s$ with $e_i > 0$ and every $1 \leq j \leq e_i$, $\mathbf{v}_{e_1 + \dots + e_i - e_i + j} = \mathbf{w}_{(i,j)}$.

If $\mathcal{B} = \emptyset$, it is vacuously linearly independent. Thus assume $e > 0$. Let (c_1, \dots, c_e) be a linear relation among \mathcal{B} . For every $i = 1, \dots, s$ with $e_i > 0$ and every $1 \leq j \leq e_i$, define $c_{(i,j)} = c_{e_1 + \dots + e_i - e_i + j}$. If $e_i = 0$, define $\mathbf{w}_i = \mathbf{0}$. If $e_i > 0$, define $\mathbf{w}_i = c_{(i,1)}\mathbf{w}_{(i,1)} + \dots + c_{(i,e_i)}\mathbf{w}_{(i,e_i)}$. Then,

$$\mathbf{0} = \sum_{k=1}^e c_k \mathbf{w}_k = \sum_{1 \leq i \leq s, e_i > 0} \left(\sum_{j=1}^{e_i} c_{e_1 + \dots + e_i - e_i + j} \mathbf{w}_{e_1 + \dots + e_i - e_i + j} \right) = \sum_{i=1}^s \mathbf{w}_i.$$

For every $i = 1, \dots, s$, because $W_i \subset V$ is a vector subspace, $\mathbf{w}_i \in W_i$, i.e., $(\mathbf{w}_1, \dots, \mathbf{w}_s)$ is an ordered s -tuple of vectors in \mathcal{W} . Because \mathcal{W} is linearly independent and because $\mathbf{w}_1 + \dots + \mathbf{w}_s = \mathbf{0}$, $\mathbf{w}_1 = \dots = \mathbf{w}_s = \mathbf{0}$. So for every $i = 1, \dots, s$ with $e_i > 0$, $(c_{(i,1)}, \dots, c_{(i,e_i)})$ is a linear relation among \mathcal{B}_i . By hypothesis, \mathcal{B}_i is linearly independent. Therefore $c_{(i,1)} = \dots = c_{(i,e_i)} = 0$. So for every $i = 1, \dots, s$ such that $e_i > 0$, and for every $1 \leq j \leq e_i$, $c_{(i,j)} = 0$, i.e., the linear relation (c_1, \dots, c_e) is $(0, \dots, 0)$. Since the only linear relation among \mathcal{B} is the trivial linear relation, it is linearly independent.

(\Leftarrow) Assume that for every ordered s -tuple $(\mathcal{B}_1, \dots, \mathcal{B}_s)$ of linearly independent ordered subset $\mathcal{B}_i \subset W_i$, the concatenation $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ is a linearly independent ordered subset of V . Let $(\mathbf{w}_1, \dots, \mathbf{w}_s)$ be an ordered s -tuple of elements in \mathcal{W} . For every $i = 1, \dots, s$, if $\mathbf{w}_i = \mathbf{0}$, define $\mathcal{B}_i = \emptyset$, otherwise define $\mathcal{B}_i = (\mathbf{w}_i)$. Each $\mathcal{B}_i \subset W_i$ is a linearly independent subset. By construction, $\mathcal{B} = (\mathbf{w}_i | 1 \leq i \leq s, \mathbf{w}_i \neq \mathbf{0})$. By hypothesis, this is linearly independent. If \mathcal{B} is nonempty, then,

$$\mathbf{0} = \sum_{i=1}^s \mathbf{w}_i = \sum_{\mathbf{w}_i \in \mathcal{B}} \mathbf{w}_i.$$

Therefore $(1, \dots, 1)$ is a nontrivial linear relation among \mathcal{B} . This contradiction proves that \mathcal{B} is empty, i.e., $\mathbf{w}_1 = \dots = \mathbf{w}_s = \mathbf{0}$. Therefore \mathcal{W} is linearly independent. \square

Proposition 1.4. *An ordered s -tuple of subspaces of V , $\mathcal{W} = (W_1, \dots, W_s)$, spans V iff for every ordered s -tuple $(\mathcal{B}_1, \dots, \mathcal{B}_s)$ of spanning subsets $\mathcal{B}_i \subset W_i$, the union $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ is a spanning set for V .*

Proof. (\Rightarrow) Assume \mathcal{W} spans V . Let $\mathbf{v} \in V$ be any vector. If $\mathbf{v} = \mathbf{0}$, by convention it is in $\text{span}(\mathcal{B})$ (even if $\mathcal{B} = \emptyset$). Assume $\mathbf{v} \neq \mathbf{0}$. Because \mathcal{W} is spanning, there exists an ordered s -tuple of vectors in \mathcal{W} , $(\mathbf{w}_1, \dots, \mathbf{w}_s)$, such that $\mathbf{v} = \mathbf{w}_1 + \dots + \mathbf{w}_s$. For every $i = 1, \dots, s$ with $\mathbf{w}_i \neq \mathbf{0}$, because \mathcal{B}_i is a spanning set for W_i , there exists $e_i > 0$, vectors $\mathbf{w}^{(i,1)}, \dots, \mathbf{w}^{(i,e_i)} \in \mathcal{B}_i$ and scalars $c_{(i,1)}, \dots, c_{(i,e_i)} \in \mathbb{F}$ such that $\mathbf{w}_i = c_{(i,1)}\mathbf{w}^{(i,1)} + \dots + c_{(i,e_i)}\mathbf{w}^{(i,e_i)}$. Then the following collection of vectors is a collection in \mathcal{B} ,

$$(\mathbf{w}^{(i,j)} | 1 \leq i \leq s, \mathbf{w}_i \neq \mathbf{0}, 1 \leq j \leq e_i).$$

For the choice of coefficients,

$$(c_{(i,j)} | 1 \leq i \leq s, \mathbf{w}_i \neq \mathbf{0}, 1 \leq j \leq e_i),$$

the linear combination of these vectors in \mathcal{B} is,

$$\sum 1 \leq i \leq s, \mathbf{w}_i \neq \mathbf{0} \mathbf{w}_i = \mathbf{v}.$$

So $\mathbf{v} \in \text{span}(\mathcal{B})$.

(\Leftarrow) Assume that for every ordered s -tuple $(\mathcal{B}_1, \dots, \mathcal{B}_s)$ of spanning subsets $\mathcal{B}_i \subset W_i$, the union \mathcal{B} is a spanning set for V . For $i = 1, \dots, s$, define $\mathcal{B}_i = W_i$. This is certainly a spanning subset of W_i . By hypothesis, $\mathcal{B} = W_1 \cup \dots \cup W_s$ is a spanning set for V .

Let $\mathbf{v} \in V$ be a vector. If $\mathbf{v} = \mathbf{0}$, then $(\mathbf{0}, \dots, \mathbf{0})$ is an ordered s -tuple of vectors in \mathcal{W} such that $\mathbf{v} = \mathbf{0} + \dots + \mathbf{0}$. Therefore assume $\mathbf{v} \neq \mathbf{0}$. Because \mathcal{B} is a spanning

set for V , exists an integer $e > 0$, nonzero vectors $\mathbf{v}_1, \dots, \mathbf{v}_e \in \mathcal{B}$, and scalars $c_1, \dots, c_e \in \mathbb{F}$ such that,

$$\mathbf{v} = c_1 \mathbf{v}_1 + \dots + c_e \mathbf{v}_e.$$

For every $j = 1, \dots, e$, let $1 \leq i(j) \leq s$ be an integer such that $\mathbf{v}_j \in W_{i(j)}$; because $\mathcal{B} = W_1 \cup \dots \cup W_s$, there is at least one such $i(j)$. For every i , if $i(j) \neq i$ for every $j = 1, \dots, e$, define $\mathbf{w}_i = \mathbf{0}$. Otherwise define,

$$\mathbf{w}_i = \sum_{1 \leq j \leq e, i(j)=i} c_j \mathbf{v}_j.$$

For every $i = 1, \dots, s$, because $W_i \subset V$ is a vector subspace, $\mathbf{w}_i \in W_i$. Thus $(\mathbf{w}_1, \dots, \mathbf{w}_s)$ is an ordered s -tuple of vectors in \mathcal{W} . And,

$$\mathbf{v} = \sum_{j=1}^e c_j \mathbf{v}_j = \sum_{i=1}^s \left(\sum_{j, i(j)=i} c_j \mathbf{v}_j \right) = \sum_{i=1}^s \mathbf{w}_i.$$

Therefore \mathcal{W} spans V . □

Proposition 1.5. *An ordered s -tuple of subspaces of V , $\mathcal{W} = (W_1, \dots, W_s)$, is a direct sum decomposition of V iff for every ordered s -tuple $(\mathcal{B}_1, \dots, \mathcal{B}_s)$ of bases \mathcal{B}_i for W_i , the union $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ is a basis for V .*

Proof. This follows from Proposition 1.3 and Proposition 1.4. □

2. GENERALIZED EIGENSPACES

Definition 2.1. Let $T : V \rightarrow V$ be a linear operator. For every integer $n \geq 0$, the n^{th} iterate of T is the linear operator $T^n : V \rightarrow V$ recursively defined by $T^0 = \text{Id}_V$, and $T^{n+1} = T \circ T^n$. For every polynomial $f(x) = a_n x^n + \dots + a_1 x + a_0 \in \mathbb{F}[x]$, the associated linear operator, $f(T) : V \rightarrow V$, is defined to be $f(T) = a_n T^n + \dots + a_1 T + a_0 \text{Id}_V$.

Proposition 2.2.

- (i) For every pair of polynomials $f(x), g(x) \in \mathbb{F}[x]$, $f(T) + g(T) = (f + g)(T)$.
- (ii) For every polynomial $f(x) \in \mathbb{F}[x]$ and every scalar $a \in \mathbb{F}$, $a \cdot (f(T)) = (a \cdot f)(T)$.
- (iii) For every pair of polynomials $f(x), g(x) \in \mathbb{F}[x]$, $f(T) \circ g(T) = (f \cdot g)(T) = g(T) \circ f(T)$.

Proof. (i) Let $f(x) = a_n x^n + \dots + a_1 x + a_0$ and let $g(x) = b_n x^n + \dots + b_1 x + b_0$. By definition of polynomial addition, $f + g = (a_n + b_n)x^n + \dots + (a_1 + b_1)x + (a_0 + b_0)$. By definition of the associated linear operator,

$$(f + g)(T) = (a_n + b_n)T^n + \dots + (a_1 + b_1)T + (a_0 + b_0)\text{Id}_V.$$

By associativity and commutativity of addition of operators, and by distributivity of scalar multiplication of operators and addition of operators, this equals,

$$(a_n T^n + \dots + a_1 T + a_0 \text{Id}_V) + (b_n T^n + \dots + b_1 T + b_0 \text{Id}_V).$$

By definition of the associated linear operator, this is $f(T) + g(T)$, i.e., $(f + g)(T) = f(T) + g(T)$.

(ii) Let $f(x) = a_n x^n + \dots + a_1 x + a_0$. By definition of scalar product of polynomials, $a \cdot f(x) = (aa_n)x^n + \dots + (aa_1)x + (aa_0)$. By definition of the associated linear operator,

$$(a \cdot f)(T) = (aa_n)T^n + \dots + (aa_1)T + (aa_0)\text{Id}_V.$$

By distributivity of multiplication of scalars and scalar multiplication of operators, this is,

$$a \cdot (a_n T^n) + \dots a \cdot (a_1 T) + a \cdot (a_0 \text{Id}_V).$$

By distributivity of scalar multiplication of operators and operator addition, this is,

$$a \cdot (a_n T^n + \dots + a_1 T + a_0 \text{Id}_V).$$

By definition of the associated linear operator, this is $a \cdot (f(T))$, i.e., $(a \cdot f)(T) = a \cdot (f(T))$.

(iii) The claim is that for every pair of integers $m, n \geq 0$, $T^m \circ T^n = T^{m+n}$. This is proved by induction on m . For $m = 0$, $T^0 = \text{Id}_V$ so that $T^0 \circ T^n = \text{Id}_V \circ T^n = T^n = T^{0+n}$. Thus assume $m > 0$ and assume the result is true for smaller values of m . By the recursive definition, $T^m \circ T^n = (T \circ T^{m-1}) \circ T^n$. By associativity of composition of linear transformations, $(T \circ T^{m-1}) \circ T^n = T \circ (T^{m-1} \circ T^n)$. By the induction hypothesis, $T^{m-1} \circ T^n = T^{m-1+n}$. Therefore $T^m \circ T^n = T \circ T^{m-1+n}$. By the recursive definition, $T^{m+n} = T \circ T^{m-1+n}$, i.e., $T^m \circ T^n = T^{m+n}$. So the claim is proved by induction on m . In particular, since $m + n = n + m$, $T^m \circ T^n = T^{m+n} = T^{n+m} = T^n \circ T^m$.

Let $f(x) = a_m x^m + \dots + a_1 x + a_0$ and let $g(x) = b_n x^n + \dots + b_1 x + b_0$. Let $N = \max(m, n)$. For $i > m$, define $a_i = 0$. For $j > n$, define $b_j = 0$. By definition of polynomial multiplication,

$$(f \cdot g)(x) = \sum_{k=0}^N \left(\sum_{i=0}^k a_i b_{k-i} \right) x^k.$$

By definition of the associated linear operator,

$$(f \cdot g)(T) = \sum_{k=0}^n \left(\sum_{i=0}^k a_i b_{k-i} \right) T^k.$$

By the claim above, $T^k = T^i \circ T^{k-i}$. Together with commutativity, associativity and distributivity, the formula above is,

$$\sum_{k=0}^n \left(\sum_{i=0}^k a_i b_{k-i} T^i \circ T^{k-i} \right) = \sum_{k=0}^n \left(\sum_{i=0}^k (a_i T^i) \circ (b_{k-i} T^{k-i}) \right).$$

By distributivity of addition and composition of linear transformations, this is,

$$\left(\sum_{i=0}^m a_i T^i \right) \circ \left(\sum_{j=0}^n b_j T^j \right).$$

By the definition of the associated linear operator, this is $f(T) \circ g(T)$, i.e., $(f \cdot g)(T) = f(T) \circ g(T)$. \square

Definition 2.3. For every linear operator $T : V \rightarrow V$ and every polynomial $f(x) \in \mathbb{F}[x]$, the *associated generalized eigenspace* is $E_{T,f} = \ker(f(T))$. For every $\lambda \in \mathbb{F}$ and every integer $n \geq 0$, the n^{th} *generalized λ -eigenspace* is $E_{T,\lambda}^{(n)} = E_{T, (x-\lambda)^n}$.

Proposition 2.4. (i) For every pair of polynomials $f(x), g(x) \in \mathbb{F}[x]$, $E_{T,f} \cap E_{T,g} \subset E_{T,f+g}$.

(ii) For every polynomial $f(x) \in \mathbb{F}[x]$ and every scalar $a \in \mathbb{F}$, $E_{T,f} \subset E_{T,a \cdot f}$. If $a \neq 0$, then $E_{T,f} = E_{T,a \cdot f}$.

(iii) For every pair of polynomials $f(x), g(x) \in \mathbb{F}[x]$, $E_{T,f} + E_{T,g} \subset E_{T,f \cdot g}$.

Proof. (i) For every $\mathbf{v} \in E_{T,f} \cap E_{T,g}$, $f(T)(\mathbf{v}) = g(T)(\mathbf{v}) = \mathbf{0}$. By Proposition 2.2(i), $(f + g)(T)(\mathbf{v}) = f(T)(\mathbf{v}) + g(T)(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Therefore $\mathbf{v} \in E_{T,f+g}$, i.e., $E_{T,f} \cap E_{T,g} \subset E_{T,f+g}$.

(ii) For every $\mathbf{v} \in E_{T,f}$, $f(T)(\mathbf{v}) = \mathbf{0}$. By Proposition 2.2(ii),

$$(a \cdot f)(T)(\mathbf{v}) = a \cdot (f(T)(\mathbf{v})) = a \cdot \mathbf{0} = \mathbf{0}.$$

Therefore $\mathbf{v} \in E_{T,a \cdot f}$, i.e., $E_{T,f} \subset E_{T,a \cdot f}$. If $a \neq 0$, then $f = a^{-1} \cdot (a \cdot f)$ so that also $E_{T,f} \subset E_{T,a \cdot f}$, i.e., $E_{T,f} = E_{T,a \cdot f}$.

(iii) For every $\mathbf{v} \in E_{T,f}$, $f(T)(\mathbf{v}) = \mathbf{0}$. By Proposition 2.2(iii),

$$(f \cdot g)(T)(\mathbf{v}) = (g \cdot f)(T)(\mathbf{v}) = (g(T) \circ f(T))(\mathbf{v}) = g(T)(f(T)(\mathbf{v})) = g(T)(\mathbf{0}) = \mathbf{0}.$$

So $\mathbf{v} \in E_{T,f \cdot g}$, i.e., $E_{T,f} \subset E_{T,f \cdot g}$. By the same argument, $E_{T,g} \subset E_{T,f \cdot g}$. Because $E_{T,f \cdot g}$ is a vector subspace, $E_{T,f} + E_{T,g} \subset E_{T,f \cdot g}$. \square

Corollary 2.5. For every $\lambda \in \mathbb{F}$, $\{\mathbf{0}\} = E_{T,\lambda}^{(0)} \subset E_{T,\lambda}^{(1)} \subset E_{T,\lambda}^{(2)} \subset \dots \subset V$.

Proof. For every integer n , by Proposition 2.4(iii), $E_{T,\lambda}^{(n)} = E_{T,(x-\lambda)^n} \subset E_{T,(x-\lambda)^{n+1}} = E_{T,\lambda}^{(n+1)}$. \square

Definition 2.6. For every $\lambda \in \mathbb{F}$, the *generalized λ -eigenspace* is $E_{T,\lambda}^{(\infty)} = \cup_{n \geq 0} E_{T,\lambda}^{(n)}$.

Proposition 2.7. For every $\lambda \in \mathbb{F}$, the *generalized λ -eigenspace* $E_{T,\lambda}^{(\infty)}$ is a vector subspace of V .

Proof. Because $\{\mathbf{0}\} = E_{T,\lambda}^{(0)} \subset E_{T,\lambda}^{(\infty)}$, the subset $E_{T,\lambda}^{(\infty)}$ is nonempty. For every $\mathbf{v}, \mathbf{w} \in E_{T,\lambda}^{(\infty)}$, there exist integers $m, n \geq 0$ such that $\mathbf{v} \in E_{T,\lambda}^{(m)}$ and $\mathbf{w} \in E_{T,\lambda}^{(n)}$. Let $N = \max(m, n)$. By Corollary 2.5, $E_{T,\lambda}^{(m)}, E_{T,\lambda}^{(n)} \subset E_{T,\lambda}^{(N)}$. Since $E_{T,\lambda}^{(N)}$ is a vector subspace, $\mathbf{v} + \mathbf{w} \in E_{T,\lambda}^{(N)} \subset E_{T,\lambda}^{(\infty)}$, i.e., $E_{T,\lambda}^{(\infty)}$ is stable under addition of elements. By a similar argument, $E_{T,\lambda}^{(\infty)}$ is stable under scalar multiplication of elements. So $E_{T,\lambda}^{(\infty)}$ is a vector subspace of V . \square

3. THE CHINESE REMAINDER THEOREM

Ancient Chinese generals discovered a beautiful and efficient method for counting large numbers of soldiers very quickly – a task of great importance in determining the number of losses after a battle. Let n_1, \dots, n_s be integers such that for every $1 \leq i < j \leq s$, the pair of integers (n_i, n_j) have no common factor. For every $i = 1, \dots, s$ have the N soldiers line up in rows of size n_i . *Do not count the number of rows!* Instead, count the remainder r_i , i.e., the number of soldiers who cannot line up in a row of size n_i . There exist integers a_1, \dots, a_s depending only on n_1, \dots, n_s such that for every integer N , N is congruent to $a_1 r_1 + \dots + a_s r_s$, modulo $n_1 \cdots n_s$. If $n_1 = 10, n_2 = 11, n_3 = 13$, the numbers are $a_1 = -429, a_2 = 650, a_3 = -220$, so that $N = -429r_1 + 650r_2 - 200r_3$, modulo 1430. These may seem like large numbers, but it is much faster for a mathematician to compute the product on an abacus than to count the soldiers by brute force. Since the general presumably knows the number of soldiers to within a range of ± 715 , this method allows the general to compute precisely the number of soldiers very quickly. We will use the same

method to prove that for distinct $\lambda_1, \dots, \lambda_s \in \mathbb{F}$, the ordered s -tuple of generalized eigenspaces, $\mathcal{W} = (E_{T, \lambda_1}^{(\infty)}, \dots, E_{T, \lambda_s}^{(\infty)})$, is linearly independent.

Definition 3.1. For an s -tuple of polynomials in $\mathbb{F}[x]$, (f_1, \dots, f_s) , not all zero, a *greatest common factor* is a polynomial $f(x)$ of maximal degree such that $f(x)$ divides $f_i(x)$ for every $i = 1, \dots, s$. If $f_1 = \dots = f_s = 0$, the greatest common factor is defined to be 0. An s -tuple of polynomials (f_1, \dots, f_s) is *coprime* if there exist polynomials (g_1, \dots, g_s) such that $1 = g_1 \cdot f_1 + \dots + g_s \cdot f_s$.

Theorem 3.2 (The Chinese remainder theorem). *An s -tuple of polynomials (f_1, \dots, f_s) is coprime iff 1 is a greatest common factor.*

Proof. (\Rightarrow) Assume (f_1, \dots, f_s) is coprime, i.e., $1 = g_1 \cdot f_1 + \dots + g_s \cdot f_s$. Let f be any common factor of (f_1, \dots, f_s) . For every i , because f_i is divisible by f , also $g_i \cdot f_i$ is divisible by f . Because every $g_i \cdot f_i$ is divisible by f , the sum $g_1 \cdot f_1 + \dots + g_s \cdot f_s$ is divisible by f , i.e., 1 is divisible by f . But the only polynomials dividing 1 are nonzero constants. Thus 1 is a greatest common factor of (f_1, \dots, f_s) .

(\Leftarrow) Assume 1 is a greatest common factor of f_1, \dots, f_s . The claim is that (f_1, \dots, f_s) is coprime. This will be proved by induction on s . If $s = 1$, then f_1 is a greatest common factor of (f_1) . Since also 1 is a greatest common factor, $\deg(f_1) = 0$, i.e., f_1 is a scalar. By definition, the greatest common factor of (0) is 0, so f_1 is nonzero. Since \mathbb{F} is a field, there exists a scalar g_1 such that $1 = g_1 \cdot f_1$, i.e., (f_1) is coprime.

Next assume $s = 2$. If necessary, permute f_1 and f_2 so that $\deg(f_1) \geq \deg(f_2)$. By hypothesis, 1 is a greatest common factor of (f_1, f_2) . If $f_2 = 0$, then 1 is a greatest common factor of (f_1) and by the last case there exists g_1 such that $1 = g_1 \cdot f_1$. Putting $g_2 = 0$, $1 = g_1 \cdot f_1 + g_2 \cdot f_2$, i.e., (f_1, f_2) is coprime. Therefore assume $f_2 \neq 0$, which implies $f_1 \neq 0$. The claim in this case will be proved by induction on $\deg(f_2)$. If $\deg(f_2) = 0$, i.e., if f_2 is a nonzero scalar, there exists a nonzero scalar g_2 such that $1 = g_2 \cdot f_2$. Setting $g_1 = 0$, $1 = g_1 \cdot f_1 + g_2 \cdot f_2$. Thus, by way of induction, assume $\deg(f_2) > 0$ and assume the result is known for smaller values of $\deg(f_2)$.

By the division algorithm, there exist polynomials $q(x)$ and $r(x)$ with $\deg(r) < \deg(f_2)$ such that $f_1(x) = q(x)f_2(x) + r(x)$. If h is a common factor of (f_2, r) , then h is also a factor of qf_2 and so of the sum $qf_2 + r$, i.e., h is a common factor of (f_1, f_2) . Conversely, if h is a common factor of (f_1, f_2) , then h is a common factor of $-qf_2$ and so of the sum $f_1 + (-qf_2)$, i.e., h is a common factor of (f_2, r) . So the common factors of (f_1, f_2) are precisely the common factors of (f_2, r) . By hypothesis, 1 is a greatest common factor of (f_2, r) . Since $\deg(r) < \deg(f_2)$, by the induction hypothesis there exist polynomials g'_1, g'_2 such that $1 = g'_1 f_2 + g'_2 r$. Define $g_1 = g'_2$ and $g_2 = g'_1 - qg'_2$. Then,

$$g_1 \cdot f_1 + g_2 \cdot f_2 = g'_2(qf_2 + r) + (g'_1 - qg'_2)f_2 = g'_1 f_2 + g'_2 r = 1.$$

So (f_1, f_2) is coprime. So the claim is proved by induction on $\deg(f_2)$ if $s = 2$.

By way of induction, assume $s > 2$ and the result is known for smaller s . Let h be a greatest common factor of (f_1, \dots, f_{s-1}) . Then 1 is a greatest common factor of $(f_1/h, \dots, f_{s-1}/h)$. By the induction hypothesis, there exist elements g'_1, \dots, g'_{s-1} such that,

$$1 = g'_1 \cdot (f_1/h) + \dots + g'_{s-1} \cdot (f_{s-1}/h),$$

i.e.,

$$h = g'_1 \cdot f_1 + \cdots + g'_{s-1} f_{s-1}.$$

The greatest common factor of $(f_1, \dots, f_{s-1}, f_s)$ is the greatest common factor of (h, f_s) . By hypothesis, 1 is a greatest common factor of (h, f_s) . By the case $s = 2$ above, there exist elements g''_1, g''_2 such that $1 = g''_1 h + g''_2 f_s$. Defining $g_1 = g''_1 g'_1, g_2 = g''_1 g'_2, \dots, g_{s-1} = g''_1 g'_{s-1}$ and defining $g_s = g''_2$,

$$g_1 f_1 + \cdots + g_{s-1} f_{s-1} + g_s f_s = g''_1 (g'_1 f_1 + \cdots + g'_{s-1} f_{s-1}) + g''_2 f_s = g''_1 h + g''_2 f_s = 1.$$

So (f_1, \dots, f_s) is coprime. The theorem is proved by induction on s . \square

Corollary 3.3. *Let $s \geq 2$ and let (f_1, \dots, f_s) be nonzero polynomials. For every $i = 1, \dots, s$, define,*

$$r_i(x) = \prod_{1 \leq j \leq s, j \neq i} f_j(x).$$

If for every $1 \leq i < j \leq s$, 1 is a greatest common factor of (f_i, f_j) , then (r_1, \dots, r_s) is coprime.

Proof. By Theorem 3.2, it suffices to prove that 1 is a greatest common factor of (r_1, \dots, r_s) . This will be proved by induction on s . For $s = 2$, $(r_1, r_2) = (f_2, f_1)$. By hypothesis, 1 is a greatest common factor. Thus, by way of induction, assume $s > 2$ and the result is known for smaller values of s . Consider the sequence (f_1, \dots, f_{s-1}) . The hypothesis holds for this collection. For every $i = 1, \dots, s-1$, denote

$$r'_i = \prod_{1 \leq j \leq s-1, j \neq i} f_j.$$

By the induction hypothesis, 1 is a common factor of (r'_1, \dots, r'_{s-1}) . Therefore f_s is a greatest common factor of $(f_s r'_1, \dots, f_s r'_{s-1}) = (r_1, \dots, r_{s-1})$. So the common factors of $(r_1, \dots, r_{s-1}, r_s)$ are the common factors of $(f_s, r_s) = (f_s, f_1 \cdots f_{s-1})$. Let h be an irreducible common factor of $(f_s, f_1 \cdots f_{s-1})$. Because h is irreducible and divides $f_1 \cdots f_{s-1}$, there exists $1 \leq i \leq s-1$ such that h divides f_i , i.e., h is a common factor of (f_i, f_s) . By hypothesis, 1 is a greatest common factor of (f_i, f_s) . Therefore h is a nonzero scalar. So for every common factor h of $(f_s, f_1 \cdots f_{s-1})$, every irreducible factor of h is a nonzero scalar, i.e., h is a nonzero scalar. Therefore 1 is a greatest common factor of (r_1, \dots, r_s) . The corollary is proved by induction on s . \square

Corollary 3.4. *For every s -tuple of distinct scalars $\lambda_1, \dots, \lambda_s \in \mathbb{F}$ and every integer $n > 0$, for every $i = 1, \dots, s$ define,*

$$r_i(x) = \prod_{1 \leq j \leq s, j \neq i} (x - \lambda_j)^n.$$

Then (r_1, \dots, r_s) is coprime.

Proof. For every $1 \leq i < j \leq n$, 1 is a greatest common factor of $(x - \lambda_i)^n$ and $(x - \lambda_j)^n$. To see this, first observe,

$$1 = (x - \lambda_i)/(\lambda_j - \lambda_i) + (x - \lambda_j)/(\lambda_i - \lambda_j).$$

Raising both sides to the power $2n - 1$ and using the binomial theorem gives,

$$1 = 1^n = \left(\sum_{i=1}^{n-1} \binom{2n-1}{i} (-1)^i (x - \lambda_i)^{n-1-i} (x - \lambda_j)^i / (\lambda_j - \lambda_i)^{2n-1} \right) (x - \lambda_i)^n + \left(\sum_{i=n}^{2n-1} \binom{2n-1}{i} (-1)^i (x - \lambda_i)^{2n-1-i} (x - \lambda_j)^{i-n} / (\lambda_j - \lambda_i)^{2n-1} \right) (x - \lambda_j)^n.$$

In other words, $((x - \lambda_i)^n, (x - \lambda_j)^n)$ is coprime. By the easy part of Theorem 3.2, 1 is a greatest common factor. \square

4. LINEAR INDEPENDENCE OF GENERALIZED EIGENSPACES

Let $s \geq 2$ be an integer. Let $\lambda_1, \dots, \lambda_s \in \mathbb{F}$ be distinct, and let $n > 0$ be an integer. Define r_1, \dots, r_s as in Corollary 3.4. By Corollary 3.4, there exist polynomials g_1, \dots, g_s such that,

$$1 = g_1(x)r_1(x) + \dots + g_s(x)r_s(x).$$

Let $T : V \rightarrow V$ be a linear operator. The identity above gives an identity of associated linear operators,

$$\text{Id}_V = 1(T) = g_1(T) \circ r_1(T) + \dots + g_s(T) \circ r_s(T).$$

Lemma 4.1. *For every $1 \leq i \leq s$ and every $\mathbf{w}_i \in E_{T, \lambda_i}^{(n)}$, for every $1 \leq j \leq s$ with $j \neq i$, $r_j(T)(\mathbf{w}_i) = \mathbf{0}$.*

Proof. Because $j \neq i$, by construction $r_j(x) = q(x)(x - \lambda_i)^n$ for some $q(x) \in \mathbb{F}[x]$, specifically,

$$q(x) = \prod_{1 \leq k \leq s, k \neq i, j} (x - \lambda_k)^n.$$

By Proposition 2.2(iii), $r_j(T) = q(T) \circ (T - \lambda_i \text{Id}_V)^n$. By hypothesis, $\mathbf{w}_i \in E_{T, \lambda_i}^{(n)} := \ker((T - \lambda_i \text{Id}_V)^n)$. Thus $r_j(T)(\mathbf{w}_i) = q(T)((T - \lambda_i \text{Id}_V)^n(\mathbf{w}_i)) = q(T)(\mathbf{0}) = \mathbf{0}$. \square

Lemma 4.2. *For every $i = 1, \dots, s$ and every $\mathbf{w}_i \in E_{T, \lambda_i}^{(n)}$, $g_i(T) \circ r_i(T)(\mathbf{w}_i) = \mathbf{w}_i$.*

Proof. By the identity, $\mathbf{w}_i = \text{Id}_V(\mathbf{w}_i) = \sum_{j=1}^s g_j(T) \circ r_j(T)(\mathbf{w}_i)$. By Lemma 4.1, for every $j \neq i$, $g_j(T) \circ r_j(T)(\mathbf{w}_i) = g_j(T)(r_j(T)(\mathbf{w}_i)) = g_j(T)(\mathbf{0}) = \mathbf{0}$. \square

Denote by \mathcal{W} the ordered s -tuple of vector subspaces of V , $\mathcal{W} = (E_{T, \lambda_1}^{(n)}, \dots, E_{T, \lambda_s}^{(n)})$

Proposition 4.3. *For every ordered s -tuple of vectors in \mathcal{W} , $(\mathbf{w}_1, \dots, \mathbf{w}_s)$, denoting $\mathbf{v} = \mathbf{w}_1 + \dots + \mathbf{w}_s$, for every $i = 1, \dots, s$, $\mathbf{w}_i = g_i(T) \circ r_i(T)(\mathbf{v})$.*

Proof. Because $g_i(T) \circ r_i(T)$ is a linear operator,

$$g_i(T) \circ r_i(T)(\mathbf{v}) = g_i(T) \circ r_i(T)\left(\sum_{j=1}^s \mathbf{w}_j\right) = \sum_{j=1}^s g_i(T) \circ r_i(T)(\mathbf{w}_j).$$

By Lemma 4.1, if $j \neq i$ then $g_i(T) \circ r_i(T)(\mathbf{w}_j) = g_i(T)(r_i(T)(\mathbf{w}_j)) = g_i(T)(\mathbf{0}) = \mathbf{0}$. Therefore,

$$g_i(T) \circ r_i(T)(\mathbf{v}) = g_i(T) \circ r_i(T)(\mathbf{w}_i).$$

By Lemma 4.2, $g_i(T) \circ r_i(T)(\mathbf{w}_i) = \mathbf{w}_i$. Therefore $g_i(T) \circ r_i(T)(\mathbf{v}) = \mathbf{w}_i$. \square

Theorem 4.4. *Let $T : V \rightarrow V$ be a linear operator. For every integer $s \geq 1$ and every ordered s -tuple of distinct scalars $(\lambda_1, \dots, \lambda_s)$, the ordered s -tuple of vector subspaces $\mathcal{W} = (E_{T, \lambda_1}^{(\infty)}, \dots, E_{T, \lambda_s}^{(\infty)})$ is linearly independent.*

Proof. If $s = 1$, this is trivial: for any subspace W of V , (W) is linearly independent. Thus assume $s \geq 2$. Let $(\mathbf{w}_1, \dots, \mathbf{w}_s)$ be an ordered s -tuple of vectors in \mathcal{W} such that $\mathbf{0} = \mathbf{w}_1 + \dots + \mathbf{w}_s$. By definition, for every $i = 1, \dots, s$, $E_{T, \lambda_i}^{(\infty)} = \cup_{n \geq 0} E_{T, \lambda_i}^{(n)}$. So for every $i = 1, \dots, s$, there exists an integer $n_i > 0$ such that $\mathbf{w}_i \in E_{T, \lambda_i}^{(n_i)}$. Define $n = \max(n_1, \dots, n_s)$. By Corollary 2.5, for every $i = 1, \dots, s$, $\mathbf{w}_i \in E_{T, \lambda_i}^{(n)}$. By Proposition 4.3, for every $i = 1, \dots, s$,

$$\mathbf{w}_i = g_i(T) \circ r_i(T)(\mathbf{0}) = \mathbf{0}.$$

Since $(\mathbf{w}_1, \dots, \mathbf{w}_s) = (\mathbf{0}, \dots, \mathbf{0})$, \mathcal{W} is linearly independent. \square